

# Matrix Completion via Max-Norm Constrained Optimization

T. Tony Cai<sup>\*</sup> and Wenxin Zhou<sup>†</sup>

## Abstract

This paper studies matrix completion under a general sampling model using the max-norm as a convex relaxation for the rank of the matrix. The optimal rate of convergence is established for the Frobenius norm loss. It is shown that the max-norm constrained minimization method is rate-optimal and it yields a more stable approximate recovery guarantee, with respect to the sampling distributions, than previously used trace-norm based approaches. The computational effectiveness of this method is also studied, based on a first-order algorithm for solving convex programs involving a max-norm constraint.

**Keywords:** Compressed sensing, low-rank matrix, matrix completion, max-norm constrained minimization, optimal rate of convergence, sparsity.

---

<sup>\*</sup>Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104. The research of Tony Cai was supported in part by NSF FRG Grant DMS-0854973, NSF Grant DMS-1208982, and NIH Grant R01 CA 127334-05.

<sup>†</sup>Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong.

# 1 Introduction

The problem of recovering a low-rank matrix from a subset of its entries, also known as *matrix completion*, has been an active topic of recent research with a range of applications including collaborative filtering (the Netflix problem) [15], multi-task learning [1], system identification [29], and sensor localization [3, 38, 39], among many others. We refer to [6] for a discussion of the above mentioned applications. Another noteworthy example is the structure-from-motion problem [10, 43] in computer vision. Let  $f$  and  $d$  be the number of frames and feature points respectively. The data are stacked into a low-rank matrix of trajectories, say  $M \in \mathbb{R}^{2f \times d}$ , such that every element of  $M$  corresponds to an image coordinate from a feature point of a rigid moving object at a given frame. Due to objects occlusions, errors on the tracking or variable out of range (i.e. images beyond the camera field of view), missing data are inevitable in real-life applications and are represented as empty entries in the matrix. Therefore, accurate and effective matrix completion methods are required, which fill in missing entries by suitable estimates.

As a direct search for the lowest-rank matrix satisfying the equality constraints is NP-hard [5], most previous work on matrix completion has focused on using the trace-norm, which is defined to be the sum of the singular values of the matrix, as a convex relaxation for the rank. This can be viewed as an analog to relaxing the *sparsity* of a vector to its  $\ell_1$ -norm, which has been shown to be effective both empirically and theoretically in compressed sensing. Several recent papers proved in different settings that a generic  $d \times d$  rank- $r$  matrix can be exactly and efficiently recovered from  $O(r d \text{poly} \log(d))$  randomly chosen entries [8, 9, 17, 35]. These results thus provide theoretical guarantees for the trace-norm constrained minimization method. In the case of recovering approximately low-rank matrices based on noisy observations, different types of trace-norm based estimators, which are akin to the Lasso and Dantzig selector used in sparse signal recovery, were proposed and well-studied [6, 20, 36, 24, 32, 23, 22, 21].

It is, however, unclear that the trace-norm is the best convex relaxation for the rank. A matrix  $M \in \mathbb{R}^{d_1 \times d_2}$  can be viewed as an operator mapping from  $\mathbb{R}^{d_2}$  to  $\mathbb{R}^{d_1}$ , its rank can be alternatively expressed as the smallest integer  $k$  such that the matrix  $M$  can be decomposed as  $M = UV^T$  for some  $U \in \mathbb{R}^{d_1 \times k}$  and  $V \in \mathbb{R}^{d_2 \times k}$ . In view of the matrix factorization  $M = UV^T$ , we would like  $U$  and  $V$  to have a small number of columns. The number of columns of  $U$  and  $V$  can be relaxed in a different way from the usual trace-norm by the so-called *max-norm* [28] which is defined by

$$\|M\|_{\max} = \min_{M=UV^T} \{\|U\|_{2,\infty} \|V\|_{2,\infty}\}, \quad (1.1)$$

where the infimum is over all factorizations  $M = UV^T$  with  $\|U\|_{2,\infty}$  being the operator norm of  $U : \ell_2^k \rightarrow \ell_\infty^{d_1}$  and  $\|V\|_{2,\infty}$  the operator norm of  $V : \ell_2^k \rightarrow \ell_\infty^{d_2}$  (or, equivalently,

$V^T : \ell_1^{d_2} \rightarrow \ell_2^k$ ) and  $k = 1, \dots, d_1 \wedge d_2$ . Note that  $\|U\|_{2,\infty}$  is the maximum  $\ell_2$  row norm of  $U$ . Since  $\ell_2$  is a Hilbert space, the factorization constant  $\|\cdot\|_{\max}$  indeed defines a norm on the space of operators between  $\ell_1^{d_2}$  and  $\ell_\infty^{d_1}$ .

The max-norm was recently proposed as another convex surrogate to the rank of the matrix. For collaborative filtering problems, the max-norm has been shown to be empirically superior to the trace-norm [40]. Foygel and Srebro [14] used the max-norm for matrix completion under the uniform sampling distribution. Their results are direct consequences of a recent bound on the excess risk for a smooth loss function, such as the quadratic loss, with a bounded second derivative [42].

Matrix completion has been well analyzed under the uniform sampling model, where observed entries are assumed to be sampled randomly and uniformly. In such a setting, the trace-norm regularized approach has been shown to have good theoretical and numerical performance. However, in some applications such as collaborative filtering, the uniform sampling model is unrealistic. For example, in the Netflix problem, the uniform sampling model is equivalent to assuming all users are equally likely to rate each movie and all movies are equally likely to be rated by any user. From a practical point of view, invariably some users are more active than others and some movies are more popular and thus rated more frequently. Hence, the sampling distribution is in fact non-uniform. In such a setting, Salakhutdinov and Srebro [37] showed that the standard trace-norm relaxation can behave very poorly, and suggested a weighted trace-norm regularizer, which incorporates the knowledge of true sampling distribution in its construction. Since the true sampling distribution is almost always unknown and can only be estimated based on the locations of those entries that are revealed in the sample, a commonly used method in practice is the empirically-weighted trace-norm [13].

In this paper we study matrix completion based on the noisy observations under a general sampling model using the max-norm as a convex relaxation for the rank. The rate of convergence for the max-norm constrained least squares estimator is obtained. Information-theoretical methods are used to establish a matching minimax lower bound under the general non-uniform sampling model. The minimax upper and lower bounds together yield the optimal rate of convergence for the Frobenius norm loss. It is shown that the max-norm regularized approach indeed provides a more stable approximate recovery guarantee, with respect to the sampling distributions, than previously used trace-norm based approaches. In the special case of the uniform sampling model, our results also show that the extra logarithmic factors in the results given in [42] could be avoided after a careful analysis to match the minimax lower bound with the upper bound (see Theorems 3.1 and 3.2 and the discussions in Section 3).

The max-norm constrained minimization problem is a convex program. The computa-

tional effectiveness of this method is also studied, based on a first-order algorithm developed in [26] for solving convex programs involving a max-norm constraint, which outperforms the semi-definite programming (SDP) method of Srebro, *et al.* [40]. We will show in Section 4 that the convex optimization problem can be implemented in polynomial time as a function of the sample size and the matrix dimensions.

The remainder of the paper is organized as follows. After introducing basic notation and definitions, Section 2 collects a few useful results on the max-norm, trace-norm and Rademacher complexity that will be needed in the rest of the paper. Section 3 introduces the model and the estimation procedure and then investigates the theoretical properties of the estimator. Both minimax upper and lower bounds are given. The results show that the max-norm constrained minimization method achieves the optimal rate of convergence over the parameter space. Comparison with past work is also given. Computation and implementation issues are discussed in Section 4. The proofs of the main results and key technical lemmas are given in Section 5.

## 2 Notations and Preliminaries

For any positive integer  $d$ , we use  $[d]$  to denote the collection of integers  $\{1, 2, \dots, d\}$ . For a vector  $u \in \mathbb{R}^d$  and  $0 < p < \infty$ , denote its  $\ell_p$ -norm by  $\|u\|_p = (\sum_{i=1}^d |u_i|^p)^{1/p}$ . In particular,  $\|u\|_\infty = \max_{i=1, \dots, d} |u_i|$  is the  $\ell_\infty$ -norm. For a matrix  $M = (M_{kl}) \in \mathbb{R}^{d_1 \times d_2}$ , let  $\|M\|_F = \sqrt{\sum_{k=1}^{d_1} \sum_{l=1}^{d_2} M_{kl}^2}$  be the Frobenius norm and let  $\|M\|_\infty = \max_{k,l} |M_{kl}|$  denote the elementwise  $\ell_\infty$ -norm. Given two norms  $\ell_p$  and  $\ell_q$  on  $\mathbb{R}^{d_1}$  and  $\mathbb{R}^{d_2}$  respectively, the corresponding operator norm  $\|\cdot\|_{p,q}$  of a matrix  $M \in \mathbb{R}^{d_1 \times d_2}$  is defined by  $\|M\|_{p,q} = \sup_{\|x\|_p=1} \|Mx\|_q$ . It is easy to verify that  $\|M\|_{p,q} = \|M^T\|_{q^*,p^*}$ , where  $(p, p^*)$  and  $(q, q^*)$  are conjugate pairs, i.e.  $\frac{1}{p} + \frac{1}{p^*} = 1$ . In particular,  $\|M\| = \|M\|_{2,2}$  is the spectral norm;  $\|M\|_{2,\infty} = \max_{k=1, \dots, d_1} \sqrt{\sum_{l=1}^{d_2} M_{kl}^2}$  is also the maximum row norm of  $M$ .

We collect in this section some known results on the max-norm, trace-norm and Rademacher complexity that will be used repeatedly later.

### 2.1 Max-norm and trace-norm

For a matrix  $M \in \mathbb{R}^{d_1 \times d_2}$ , its trace-norm  $\|M\|_*$  can also be equivalently written as

$$\|M\|_* = \inf \left\{ \sum_j |\sigma_j| : M = \sum_j \sigma_j u_j v_j^T, \|u_j\|_2 = \|v_j\|_2 = 1 \right\}.$$

In other words, the trace-norm promotes low-rank decompositions with factors in  $\ell_2$ . Similarly, using Grothendieck's inequality [18], the max-norm defined in (1.1) has the following

analogous representation in terms of factors in  $\ell_\infty$ :

$$\|M\|_{\max} \approx \inf \left\{ \sum_j |\sigma_j| : M = \sum_j \sigma_j u_j v_j^T, \|u_j\|_\infty = \|v_j\|_\infty = 1 \right\}.$$

The factor of equivalence is the Grothendieck's constant  $K_G \in (1.67, 1.79)$ . Based on these properties, Lee, *et al.* (2010) [26] expected max-norm regularization to be more effective when dealing with uniformly bounded data.

Of the same flavor as the definition of the max-norm in (1.1), the trace-norm has the following equivalent characterization in terms of the matrix factorization [41],

$$\|M\|_* = \min_{M=UV^T} \{\|U\|_F \|V\|_F\} = \frac{1}{2} \min_{U,V:M=UV^T} (\|U\|_F^2 + \|V\|_F^2).$$

It is easy to see that

$$\frac{\|M\|_*}{\sqrt{d_1 d_2}} \leq \|M\|_{\max}, \quad (2.1)$$

which in turn implies that any low max-norm approximation is also a low trace-norm approximation. As pointed out in [41], there can be a large gap between  $\frac{1}{\sqrt{d_1 d_2}} \|\cdot\|_*$  and  $\|\cdot\|_{\max}$ . The following relationship between the trace-norm and Frobenius norm is well-known,

$$\|M\|_F \leq \|M\|_* \leq \sqrt{\text{rank}(M)} \cdot \|M\|_F.$$

In the same spirit, an analogous bound holds for the max-norm, in connection with the element-wise  $\ell_\infty$ -norm [28]:

$$\|M\|_\infty \leq \|M\|_{\max} \leq \sqrt{\text{rank}(M)} \cdot \|M\|_{1,\infty} \leq \sqrt{\text{rank}(M)} \cdot \|M\|_\infty. \quad (2.2)$$

For any  $R > 0$ , let

$$\mathbb{B}_{\max}(R) := \{M \in \mathbb{R}^{d_1 \times d_2} : \|M\|_{\max} \leq R\} \quad \text{and} \quad \mathbb{B}_*(R) := \{M \in \mathbb{R}^{d_1 \times d_2} : \|M\|_* \leq R\} \quad (2.3)$$

be the max-norm and trace-norm ball with radius  $R$ , respectively. It is now well-known [41] that  $\mathbb{B}_{\max}(1)$  can be bounded, from both below and above, by the convex hull of rank-one sign matrices  $\mathcal{M}_\pm = \{M \in \{\pm 1\}^{d_1 \times d_2} : \text{rank}(M) = 1\}$ . That is,

$$\text{conv} \mathcal{M}_\pm \subset \mathbb{B}_{\max}(1) \subset K_G \cdot \text{conv} \mathcal{M}_\pm \quad (2.4)$$

with  $K_G \in (1.67, 1.79)$  denoting the Grothendieck's constant. Moreover,  $\mathcal{M}_\pm$  is a finite class with cardinality  $|\mathcal{M}_\pm| = 2^{d-1}$  where  $d = d_1 + d_2$ .

## 2.2 Rademacher complexity

A technical tool used in our analysis involves data-dependent estimates of the Rademacher and Gaussian complexities of a function class. We refer to Bartlett and Mendelson [2] and references therein for a detailed introduction of these concepts.

**Definition 2.1.** For a class  $\mathcal{F}$  of functions mapping from  $\mathcal{X}$  to  $\mathbb{R}$ , its empirical Rademacher complexity over a specific sample  $S = (x_1, x_2, \dots) \subset \mathcal{X}$  is given by

$$\hat{R}_S(\mathcal{F}) = \frac{2}{|S|} \mathbb{E}_\varepsilon \left[ \sup_{f \in \mathcal{F}} \left| \sum \varepsilon_i f(x_i) \right| \right], \quad (2.5)$$

where  $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots)$  is a Rademacher sequence. The Rademacher complexity with respect to a distribution  $\mathcal{P}$  is the expectation, over an i.i.d. sample of  $|S|$  points drawn from  $\mathcal{P}$ , denoted by

$$R_{|S|}(\mathcal{F}) = \mathbb{E}_{S \sim \mathcal{P}} [\hat{R}_S(\mathcal{F})].$$

Replacing  $\varepsilon_1, \dots, \varepsilon_n$  with independent Gaussian  $N(0, 1)$  random variables  $g_1, \dots, g_n$  leads to the definition of (empirical) Gaussian complexity.

Considering a matrix as a function from the index pairs to the entry values, Srebro and Shraibman [41] obtained upper bounds on the Rademacher complexity of the unit balls under both the trace-norm and the max-norm. Specifically, for any  $d_1, d_2 > 2$  and any sample of size  $2 < |S| < d_1 d_2$ , the empirical Rademacher complexity of the max-norm unit ball is bounded by

$$\hat{R}_S(\mathbb{B}_{\max}(1)) \leq 12 \sqrt{\frac{d_1 + d_2}{|S|}}. \quad (2.6)$$

## 3 Max-Norm Constrained Minimization

### 3.1 The model

We now consider matrix completion under a general random sampling model. Let  $M_0 \in \mathbb{R}^{d_1 \times d_2}$  be an unknown matrix. Suppose that a random sample

$$S = \{(i_1, j_1), (i_2, j_2), \dots, (i_n, j_n)\}$$

of the index set is drawn i.i.d. according to a general sampling distribution  $\Pi = \{\pi_{kl}\}$  on  $[d_1] \times [d_2]$ , with replacement, i.e.  $\mathbb{P}[(i_t, j_t) = (k, l)] = \pi_{kl}$  for all  $t$  and  $(k, l)$ . Given the random index subset  $S = \{(i_1, j_1), \dots, (i_n, j_n)\}$ , we observe noisy entries  $\{Y_{i_t, j_t}\}_{t=1}^n$  indexed by  $S$ , i.e.

$$Y_{i_t, j_t} = (M_0)_{i_t, j_t} + \sigma \xi_t, \quad t = 1, \dots, n, \quad (3.1)$$

for some  $\sigma > 0$ . The noise variables  $\xi_t$  are independent, with  $\mathbb{E}[\xi_t] = 0$  and  $\mathbb{E}[\xi_t^2] = 1$ .

Instead of assuming the uniform sampling distribution, we consider a general sampling distribution  $\Pi$  here. Since  $\sum_{k=1}^{d_1} \sum_{l=1}^{d_2} \pi_{kl} = 1$ , we have  $\max_{k,l} \{\pi_{kl}\} \geq \frac{1}{d_1 d_2}$ . In addition, to ensure that each entry is observed with a positive probability, it is natural to assume that there exists a positive constant  $\mu \geq 1$  such that

$$\pi_{kl} \geq \frac{1}{\mu d_1 d_2}, \quad \text{for all } (k, l) \in [d_1] \times [d_2]. \quad (3.2)$$

We write hereafter  $d = d_1 + d_2$  for brevity. Clearly,  $\max(d_1, d_2) \leq d \leq 2 \max(d_1, d_2)$ .

Past work on matrix completion has mainly focused on the case of exact low-rank matrices. Here we allow a relaxation of this assumption and consider the more general setting of approximately low-rank matrices. Specifically, we consider recovery of matrices with  $\ell_\infty$ -norm and max-norm constraints defined by

$$\mathcal{K}(\alpha, R) := \left\{ M \in \mathbb{R}^{d_1 \times d_2} : \|M\|_\infty \leq \alpha, \|M\|_{\max} \leq R \right\}. \quad (3.3)$$

Here both  $\alpha$  and  $R$  are free parameters to be determined. It is clear that  $R \geq \alpha$  is needed to guarantee that  $\mathcal{K}(\alpha, R)$  is non-empty. If  $M_0$  is of rank at most  $r$  and  $\|M_0\|_\infty \leq \alpha$ , then by (2.2) we have  $M_0 \in \mathbb{B}_{\max}(\alpha\sqrt{r})$  and hence  $M_0 \in \mathcal{K}(\alpha, \alpha\sqrt{r})$ .

### 3.2 Max-norm constrained least squares estimator

Given a collection of observations  $Y_S = \{Y_{i_t, j_t}\}_{t=1}^n$  from the observation model (3.1), we estimate the unknown  $M_0 \in \mathcal{K}(\alpha, R)$  for some  $R \geq \alpha > 0$  by the minimizer of the empirical risk with respect the quadratic loss function

$$\hat{\mathcal{L}}_n(M; Y) = \frac{1}{n} \sum_{t=1}^n (Y_{i_t, j_t} - M_{i_t, j_t})^2.$$

That is,

$$\hat{M}_{\max} := \arg \min_{M \in \mathcal{K}(\alpha, R)} \hat{\mathcal{L}}_n(M; Y). \quad (3.4)$$

The minimization procedure requires that all the entries of  $M_0$  are bounded in absolute value by a known constant  $\alpha$ . This condition enforces that  $M_0$  should not be too “spiky”, and a too large bound may jeopardize exactness of the estimation, see, e.g. [24, 32, 21]. On the other hand, as argued in Lee, *et al.* [26], the max-norm regularization is expected to be more effective particularly for uniformly bounded data, which is our main motivation for using the max-norm constrained estimator.

Although the max-norm constrained minimization problem (3.4) is a convex program, fast and efficient algorithms for solving large-scale optimization problems that incorporate the max-norm have only been developed recently [26]. We will show in Section 4 that the convex optimization problem (3.4) can indeed be implemented in polynomial time as a function of the sample size  $n$  and the matrix dimensions  $d_1$  and  $d_2$ .

### 3.3 Upper bounds

We now state our main results concerning the recovery of an approximately low-rank matrix  $M_0$  using the max-norm constrained minimization method.

**Theorem 3.1.** *Suppose that the noise sequence  $\{\xi_t\}$  are i.i.d. standard normal random variables, and the unknown matrix  $M_0 \in \mathcal{K}(\alpha, R)$  for some  $R \geq \alpha > 0$ . Then there exists an absolute constants  $C$  such that for any  $t \in (0, 1)$  and a sample size  $2 < n \leq d_1 d_2$ ,*

$$\sum_{k=1}^{d_1} \sum_{l=1}^{d_2} \pi_{kl} (\hat{M}_{\max} - M_0)_{kl}^2 \leq C \left\{ (\alpha \vee \sigma) R \sqrt{\frac{d}{n}} + \frac{\alpha^2 \log(2/t)}{n} \right\} \quad (3.5)$$

*holds with probability greater than  $1 - e^{-d} - t$ . If, in addition, the assumption (3.2) is satisfied, then for a sample size  $d \leq n \leq d_1 d_2$ ,*

$$\frac{1}{d_1 d_2} \|\hat{M}_{\max} - M_0\|_F^2 \leq C \mu (\alpha \vee \sigma) R \sqrt{\frac{d}{n}} \quad (3.6)$$

*holds with probability at least  $1 - 2e^{-d}$ .*

**Remark 3.1.** The upper bounds given in Theorem 3.1 hold with high probability. The rate of convergence under expectation can be obtained as a direct consequence. More specifically, for a sample size  $n$  with  $d \leq n \leq d_1 d_2$ , we have

$$\sup_{M_0 \in \mathcal{K}(\alpha, R)} \frac{1}{d_1 d_2} \mathbb{E} \|\hat{M}_{\max} - M_0\|_F^2 \leq C \mu (\alpha \vee \sigma) R \sqrt{\frac{d}{n}}. \quad (3.7)$$

It can be seen from the proof of Theorem 3.1 that the normality assumption on the noise can be relaxed to a class of *sub-exponential* random variables, those with at least an exponential tail decay.

**Corollary 3.1.** *Under the assumptions of Theorem 3.1, but assume instead that the noise sequence  $\{\xi_t\}$  are independent sub-exponential random variables, that is, there is a constant  $K > 0$  such that*

$$\max_{t=1, \dots, n} \mathbb{E}[\exp(|\xi_t|/K)] \leq e. \quad (3.8)$$

*Then, for a sample size  $d < n \leq d_1 d_2$ ,*

$$\sum_{k=1}^{d_1} \sum_{l=1}^{d_2} \pi_{kl} (\hat{M}_{\max} - M_0)_{kl}^2 \leq C (\alpha \vee \sigma K) R \sqrt{\frac{d}{n}}, \quad (3.9)$$

*with probability greater than  $1 - 2e^{-d}$ , where  $C > 0$  is an absolute constant.*



### 3.4 Information-theoretic lower bounds

Theorem 3.1 gives the rate of convergence for the max-norm constrained least squares estimator  $\hat{M}_{\max}$ . In this section we shall use information-theoretical methods to establish a minimax lower bound for *non-uniform sampling at random* matrix completion on the max-norm ball. The minimax lower bound matches the rate of convergence given in (3.6) when the sampling distribution  $\Pi$  satisfies  $\frac{\mu}{d_1 d_2} \leq \min_{k,l} \{\pi_{kl}\} \leq \max_{k,l} \{\pi_{kl}\} \leq \frac{L}{d_1 d_2}$  for some constants  $\mu$  and  $L$ . The results show that the max-norm constrained least squares estimator is indeed rate-optimal in such a setting.

For the lower bound, we shall assume the sampling distribution  $\Pi$  satisfies

$$\max_{k,l} \{\pi_{kl}\} \leq \frac{L}{d_1 d_2} \quad (3.10)$$

for a positive constant  $L \geq 1$ . Clearly, when  $L = 1$ , this amounts to say that the sampling distribution is uniform.

**Theorem 3.2.** *Suppose that the noise sequence  $\{\xi_t\}$  are i.i.d. standard normal random variables, the sampling distribution  $\Pi$  obeys the condition (3.10) and the quintuple  $(n, d_1, d_2, \alpha, R)$  satisfies*

$$\frac{48\alpha^2}{d_1 \vee d_2} \leq R^2 \leq \frac{\sigma^2(d_1 \wedge d_2)d_1 d_2}{128Ln}. \quad (3.11)$$

*Then the minimax  $\|\cdot\|_F$ -risk is lower bounded as*

$$\inf_{\hat{M}} \sup_{M \in \mathcal{K}(\alpha, R)} \frac{1}{d_1 d_2} \mathbb{E} \|\hat{M} - M\|_F^2 \geq \min \left\{ \frac{\alpha^2}{16}, \frac{\sigma R}{256} \sqrt{\frac{d}{nL}} \right\}. \quad (3.12)$$

*In particular, for a sample size  $n \geq (\frac{R}{\alpha})^2 \frac{d}{L}$ , we have*

$$\inf_{\hat{M}} \sup_{M \in \mathcal{K}(\alpha, R)} \frac{1}{d_1 d_2} \mathbb{E} \|\hat{M} - M\|_F^2 \geq \frac{(\alpha \wedge \sigma)R}{256} \sqrt{\frac{d}{nL}}. \quad (3.13)$$

Assume that both  $\mu$  and  $L$ , respectively appeared in (3.2) and (3.10), are bounded above by universal constants, then comparing the lower bound (3.13) with the upper bound (3.7) shows that if the sample size  $n > (R/\alpha)^2 d$ , the optimal rate of convergence is  $(\alpha \wedge \sigma)R\sqrt{\frac{d}{n}}$ , i.e.,

$$\inf_{\hat{M}} \sup_{M \in \mathcal{K}(\alpha, R)} \frac{1}{d_1 d_2} \mathbb{E} \|\hat{M} - M\|_F^2 \asymp (\alpha \wedge \sigma)R\sqrt{\frac{d}{n}},$$

and the max-norm constrained minimization estimator (3.4) is rate-optimal. The requirement here on the sample size  $n > (R/\alpha)^2 d$  is weak.

The proof of Theorem 3.2 follows the same outline as in [32], using information-theoretic methods. A key technical tool for the proof is the following lemma which guarantees the existence of a suitably large packing set for  $\mathcal{K}(\alpha, R)$  in the Frobenius norm.

**Lemma 3.1.** *Let  $r = (R/\alpha)^2$  and let  $\gamma \leq 1$  be such that  $\frac{r}{\gamma^2} \leq d_1 \wedge d_2$  is an integer. There exists a subset  $\mathcal{M} \subset \mathcal{K}(\alpha, R)$  with cardinality*

$$|\mathcal{M}| = \left\lceil \exp \left( \frac{r(d_1 \vee d_2)}{16\gamma^2} \right) \right\rceil + 1$$

*and with the following properties:*

- (i) *For any  $M \in \mathcal{M}$ ,  $\text{rank}(M) \leq \frac{r}{\gamma^2}$  and  $M_{kl} \in \{\pm\gamma\alpha\}$ , such that*

$$\|M\|_\infty = \gamma\alpha \leq 1, \quad \frac{1}{d_1 d_2} \|M\|_F^2 = \gamma^2 \alpha^2.$$

- (ii) *For any two distinct  $M^k, M^l \in \mathcal{M}$ ,*

$$\frac{1}{d_1 d_2} \|M^k - M^l\|_F^2 > \frac{\gamma^2 \alpha^2}{2}.$$

The proof of Lemma 3.1 follows from Lemma 3 of [12] with a simple modification, which for self-containment, is given in Section 5.3.

### 3.5 Comparison to past work

It is now well-known that the exact recovery of a low-rank matrix in the noiseless case requires the “incoherence conditions” on the target matrix  $M_0$  [8, 9, 35, 17]. Instead, we consider here the more general setting of approximately low-rank matrices, and prove that approximate recovery is still possible without the subtle structural conditions.

Our results are directly comparable to those of Negahban and Wainwright [32], in which the trace-norm was used as a proxy to the rank. Specifically, Negahban and Wainwright [32] considered the setting where the sampling distribution is a *product distribution*, i.e.

$$\pi_{kl} = \pi_{k\cdot} \pi_{\cdot l}, \quad \text{for all } (k, l) \in [d_1] \times [d_2],$$

where  $\pi_{k\cdot}$  and  $\pi_{\cdot l}$  are marginals satisfying

$$\pi_{k\cdot} \geq \frac{1}{\sqrt{\nu} d_1}, \quad \pi_{\cdot l} \geq \frac{1}{\sqrt{\nu} d_2} \quad \text{for some } \nu \geq 1. \quad (3.14)$$

Accordingly, define the weighted norms as

$$\|M\|_{w(\dagger)} := \|\sqrt{W_r} M \sqrt{W_c}\|_{\dagger}, \quad \dagger \in \{F, *, \infty\},$$

where  $W_r = d_1 \cdot \text{diag}(\pi_{1\cdot}, \dots, \pi_{d_1\cdot})$  and  $W_c = d_2 \cdot \text{diag}(\pi_{\cdot d_1}, \dots, \pi_{\cdot d_2})$ . Assuming that the unknown matrix  $M_0$  satisfies

$$\|M_0\|_{w(*)} \leq R\sqrt{d_1 d_2}, \quad \|M_0\|_{w(F)} \leq \sqrt{d_1 d_2} \quad \text{and} \quad \frac{\|M_0\|_{w(\infty)}}{\|M_0\|_{w(F)}} \leq \frac{\alpha}{\sqrt{d_1 d_2}},$$

then based on a collection of observations

$$Y_{i_t, j_t} = \varepsilon_t (M_0)_{i_t, j_t} + \sigma \xi_t, \quad t = 1, \dots, n$$

where  $(i_t, j_t)$  are i.i.d. according to  $\mathbb{P}[(i_t, j_t) = (k, l)] = \pi_{kl}$  and  $\varepsilon_t \in \{-1, +1\}$  are i.i.d. random signs, they proposed the following estimator of  $M_0$

$$\hat{M}_* \in \arg \min_{\|M\|_{w(\infty)} \leq \alpha} \left\{ \frac{1}{n} \sum_{t=1}^n (Y_{i_t, j_t} - \varepsilon_t M_{i_t, j_t})^2 + \lambda_n \|M\|_{w(*)} \right\} \quad (3.15)$$

and proved that for properly chosen  $\lambda_n$  depending on  $\sigma$ , there exist absolute constants  $c_i$  such that

$$\frac{1}{d_1 d_2} \|\hat{M}_* - M_0\|_F^2 \leq c_1 \nu \left\{ (\sigma \vee \nu) R \alpha \sqrt{\frac{d \log(d)}{n}} + \frac{\nu \alpha^2}{n} \right\}, \quad (3.16)$$

holds with probability at least  $1 - c_2 \exp(-c_3 \log d)$ .

First, the product distribution assumption is very restrictive and is not valid in many applications. For example, in the case of the Netflix problem, this assumption would imply that conditional on any movie, it will be rated by all users with the same probability. Second, the constraint on  $M_0$  highly depends on the true sampling distribution which is really unknown in practice and can only be estimated based on the empirical frequencies, i.e. for any pair  $(k, l) \in [d_1] \times [d_2]$ ,

$$\hat{\pi}_{k\cdot} = \frac{\sum_{t=1}^n 1_{\{i_t=k\}}}{n}, \quad \hat{\pi}_{\cdot l} = \frac{\sum_{t=1}^n 1_{\{j_t=l\}}}{n}.$$

Since only a relatively small sample of the entries of  $M_0$  is observed, these estimates are unlikely to be accurate. The max-norm constrained minimization approach, on the other hand, is proved (Theorem 3.1) to be effective in the presence of non-degenerate general sampling distributions. The method does not require either a product distribution or the knowledge of the exact true sampling distribution. Hence, the max-norm constrained method indeed yields a more robust approximate recovery guarantee, with respect to the sampling distributions.

We now turn to the special case of uniform sampling. The “spikeness” assumption in [32] can actually be reduced to a single constraint on the  $\ell_\infty$ -norm (see, e.g. [21]). Let  $\mathbb{B}_\infty(\alpha) = \{M \in \mathbb{R}^{d_1 \times d_2} : \|M\|_\infty \leq \alpha\}$  be the  $\ell_\infty$ -norm ball with radius  $\alpha$ . Define the class of matrices

$$\mathcal{K}_*(\alpha, R) := \left\{ M \in \mathbb{B}_\infty(\alpha) : \frac{\|M\|_*}{\sqrt{d_1 d_2}} \leq R \right\}. \quad (3.17)$$

It can be seen from (2.1) and (2.2) that  $\{M \in \mathbb{B}_\infty(\alpha) : \text{rank}(M) \leq r\} \subsetneq \mathcal{K}(\alpha, \alpha\sqrt{r}) \subsetneq \mathcal{K}_*(\alpha, \alpha\sqrt{r})$ . The following results provide upper bounds on the accuracy of both the max- and trace-norm regularized estimators under the Frobenius norm.

**Corollary 3.2.** *Suppose that the noise sequence  $\{\xi_t\}$  are i.i.d.  $N(0, 1)$  random variables and the sampling distribution  $\Pi$  is uniform on  $[d_1] \times [d_2]$ . Then the following inequalities hold with probability at least  $1 - 3/d$ :*

(i) *The optimum  $\hat{M}_{\max}$  to the convex program (3.4) satisfies*

$$\sup_{M_0 \in \mathcal{K}(\alpha, R)} \frac{1}{d_1 d_2} \|\hat{M}_{\max} - M_0\|_F^2 \lesssim (\sigma \vee \alpha) R \sqrt{\frac{d}{n}} + \frac{\alpha^2 \log(d)}{n}. \quad (3.18)$$

(ii) *The minima  $\hat{M}_*$  to the SDP (3.15) with all weighted norms replaced by the standard ones and with a properly chosen  $\lambda_n$  satisfies*

$$\sup_{M_0 \in \mathcal{K}_*(\alpha, R)} \frac{1}{d_1 d_2} \|\hat{M}_* - M_0\|_F^2 \lesssim (\sigma \vee \alpha) R \sqrt{\frac{d \log(d)}{n}} + \frac{\alpha^2 \log(d)}{n}. \quad (3.19)$$

The upper bound (3.18) follows immediately from (3.5) in Theorem 3.1, and (3.19) is a direct extension of Theorem 7 in Klopp (2012) which considers the case of the exact low-rank matrices, i.e.  $\text{rank}(M_0) \leq r$ . The proof is essentially the same and thus is omitted.

Foygel and Srebro [14] analyzed the estimation error of  $\hat{M}_{\max}$  based on an excess risk bound for empirical risk minimization with a smooth loss function recently developed in [42]. Specifically, assuming sub-exponential noise and  $M_0 \in \mathcal{K}(\alpha, R)$ , it was shown that with high probability,

$$\frac{1}{d_1 d_2} \|\hat{M}_{\max} - M_0\|_F^2 \lesssim (\sigma \vee \alpha) R \sqrt{\frac{d \log^3(n/d)}{n}} + \frac{R^2 d \log^3(n/d)}{n}. \quad (3.20)$$

After a more delicate analysis, our result shows that the additional  $\log^3(n/d)$  factor in (3.20) is purely an artifact of the proof technique and thus can be avoided. Moreover, in view of the lower bounds given in Theorem 3.2, we see that the max-norm constrained least square estimator  $\hat{M}_{\max}$  achieves the optimal rate of convergence for recovering approximately low-rank matrices over the parameter space  $\mathcal{K}(\alpha, R)$  under the Frobenius norm loss. To our knowledge, the best known rate for trace-norm regularized estimator ((3.19)) is near-optimal up to logarithmic factors in a minimax sense, over a larger parameter space  $\mathcal{K}_*(\alpha, R)$ .

## 4 Computational Algorithm

Although Theorem 3.1 presents theoretical guarantees that hold uniformly for any global minimizer, it does not provide guidance on how to approximate such a global minimizer using a polynomial-time algorithm. A parallel line of work has studied computationally

efficient algorithms for solving problems with the trace-norm constraint or penalization. See, for instance, Mazumber, *et al.* [30], Nesterov [33] and Lin, *et al.* [27] among others. Here we restrict our attention to the less-studied max-norm based approach. We recommend using the fast first-order algorithms developed in Lee, *et al.* [26], which is particularly tailored for large scale optimization problems that incorporate the max-norm. The problem of interest to us is the optimization program (3.4) with both the max-norm and the element-wise  $\ell_\infty$ -norm constraints, in which case the algorithm introduced in [26] can be applied only after some slight modifications as described below.

Due to Srebro, *et al.* [40], the max-norm of a  $d_1 \times d_2$  matrix  $M$  can be computed via a semi-definite program:

$$\|M\|_{\max} = \min R \quad \text{s.t.} \quad \begin{pmatrix} W_1 & M \\ M^T & W_2 \end{pmatrix} \succeq 0, \quad \text{diag}(W_1) \leq R, \quad \text{diag}(W_2) \leq R.$$

Correspondingly, we can reformulate (3.4) as the following SDP problem

$$\min f(M; Y) \quad \text{s.t.} \quad \begin{pmatrix} W_1 & M \\ M^T & W_2 \end{pmatrix} \succeq 0, \quad \text{diag}(W_1) \leq R, \quad \text{diag}(W_2) \leq R, \quad \|M\|_\infty \leq \alpha,$$

where the objective function  $f$  is given by

$$f(M; Y) = \hat{\mathcal{L}}_n(M; Y).$$

This SDP can be solved using standard interior-point methods, though are fairly slow and do not scale to matrices with large dimensions. For large-scale problems, an alternative factorization method based on (1.1), as described below, is preferred [26].

We begin by introducing dummy variables  $U \in \mathbb{R}^{d_1 \times k}$ ,  $V \in \mathbb{R}^{d_2 \times k}$  for some  $1 \leq k \leq d_1 + d_2$  and let  $M = UV^T$ . If the optimal solution  $\hat{M}_{\max}$  is known to have rank at most  $r$ , we can take  $U \in \mathbb{R}^{d_1 \times (r+1)}$ ,  $V \in \mathbb{R}^{d_2 \times (r+1)}$ . In practice, without a known guarantee on the rank of  $\hat{M}_{\max}$ , we alternatively truncate the number of columns  $k$  to some reasonably high value less than  $d_1 + d_2$ . Then we rewrite the original problem (3.4) in the factored form as follows:

$$\begin{aligned} & \text{minimize} && f(UV^T; Y) \\ & \text{subject to} && \max\{\|U\|_{2,\infty}^2, \|V\|_{2,\infty}^2\} \leq R, \quad \max_{i,j} |U_i^T V_j| \leq \alpha. \end{aligned} \quad (4.1)$$

This problem is non-convex, since it involves a constraint on all product factorizations  $UV^T$ . However, when the size of the problem (i.e.  $k$ ) is large enough, Burer and Choi (2006) proved that this reformulated problem has no local minima. To solve this problem fast and efficiently, Lee, *et al.* [26] suggest the following first-order methods.

## 4.1 Projected gradient method

Notice that  $f(M; Y) = \hat{\mathcal{L}}(M; Y)$  is differentiable with respect to the first argument. The method of projected gradient descent generates a sequence of iterates  $\{(U^t, V^t), t = 0, 1, 2, \dots\}$  by the recursion: First define an intermediate iterate

$$\begin{bmatrix} \tilde{U}^{t+1} \\ \tilde{V}^{t+1} \end{bmatrix} = \begin{bmatrix} U^t - \tau \cdot \nabla f(U^t(V^t)^T; Y)V^t \\ V^t - \tau \cdot \nabla f(U^t(V^t)^T; Y)^T U^t \end{bmatrix},$$

where  $\tau > 0$  is a stepsize parameter. If  $\|\tilde{U}^{t+1}(\tilde{V}^{t+1})^T\|_\infty > \alpha$ , we replace

$$\begin{bmatrix} \tilde{U}^{t+1} \\ \tilde{V}^{t+1} \end{bmatrix} \quad \text{with} \quad \frac{\sqrt{\alpha}}{\|\tilde{U}^{t+1}(\tilde{V}^{t+1})^T\|_\infty^{1/2}} \begin{bmatrix} \tilde{U}^{t+1} \\ \tilde{V}^{t+1} \end{bmatrix},$$

otherwise we keep it still. Next, compute updates according to

$$\begin{bmatrix} U^{t+1} \\ V^{t+1} \end{bmatrix} = \Pi_R \left( \begin{bmatrix} \tilde{U}^{t+1} \\ \tilde{V}^{t+1} \end{bmatrix} \right),$$

where  $\Pi_R$  denotes the Euclidean projection onto the set  $\{(U, V) : \max(\|U\|_{2,\infty}^2, \|V\|_{2,\infty}^2) \leq R\}$ . This projection can be computed by re-scaling the rows of the current iterate whose  $\ell_2$ -norms exceed  $R$  so that their norms become exactly  $R$ , while rows with norms already less than  $R$  remain unchanged.

## 4.2 Stepwise gradient

For the matrix completion problem, we allow the objective function to act on matrices via the average loss function over their entries:

$$f(M; Y) = f(UV^T; Y) = \frac{1}{n} \sum_{t=1}^n g(U_{i_t}^T V_{j_t}; Y_{i_t, j_t}),$$

where  $S = \{(i_1, j_1), \dots, (i_n, j_n)\} \subset [d_1] \times [d_2]$  is a training set of row-column indices,  $U_i$  and  $V_j$  denote the  $i$ th row of  $U$  and  $j$ th row of  $V$ , respectively. We are currently interested in the case where  $g(t; y) = (t - y)^2$ .

In view of the above decomposition for  $f$ , it is thus natural to use a stepwise gradient method: enumerate all elements of  $S$  in an arbitrary order with repeated ones only counted once; for the pair  $(i_t, j_t)$  at the  $t$ -th iteration, take a step in the direction opposite to the gradient of  $g(U_{i_t}^T V_{j_t}; Y_{i_t, j_t})$ , then apply the rescaling and the projection described in the last subsection. More precisely, if  $|U_{i_t}^T V_{j_t}| > \alpha$ ,  $U_{i_t}$  and  $V_{j_t}$  are replaced with  $\frac{\sqrt{\alpha} U_{i_t}}{|U_{i_t}^T V_{j_t}|^{1/2}}$  and  $\frac{\sqrt{\alpha} V_{j_t}}{|U_{i_t}^T V_{j_t}|^{1/2}}$  respectively, otherwise we do not make any change; next, if  $\|U_{i_t}\|_2^2 > R$ , we project it back so that  $\|U_{i_t}\|_2^2 = R$ , otherwise we do not make any change (the same

procedure for  $V_{j_t}$ ). In the  $t$ -th iteration, we do not need to consider any other rows of  $U$  and  $V$ . As demonstrated in [26], this stepwise algorithm could be computationally as efficient as optimization with the trace-norm.

### 4.3 Implementation

Before the max-norm constraint approach can be actually implemented in practice to generate a full matrix by filling in missing entries, additional prior knowledge of the unknown true matrix is needed to avoid deviated results. As before, let  $M_0 \in \mathbb{R}^{d_1 \times d_2}$  be the true underlying matrix. Good upper bounds for the following key quantities are needed in advance:

$$\alpha_0 = \|M_0\|_\infty, \quad R_0 = \|M_0\|_{\max} \quad \text{and} \quad r_0 = \text{rank}(M_0). \quad (4.1)$$

In order to estimate  $R_0$  directly from a missing data matrix, it can be seen from (2.2) that  $\alpha_0 \sqrt{r_0}$  is a sharp upper bound on  $R_0$  and is more amenable to estimation. Fortunately, it is possible to convincingly specify  $\alpha_0$  beforehand in many real-life applications. When dealing with the Netflix data, for instance,  $\alpha_0$  can be chosen as the highest rating index; in the structure-from-motion problem,  $\alpha_0$  depends on the range of the camera field of view, which in most cases is sufficiently large to capture the feature point trajectories. In case where the percentage of missing entries is low, the largest magnitude of the observed entries can be used as an alternative for  $\alpha_0$ .

As for  $r_0$ , we recommend the rank estimation approach recently developed in [19], which was shown to be effective in computer vision problems. Recall that in the structure-from-motion problem, each column of the data matrix corresponds a trajectory along the frames of a given feature point, and can be regarded as a signal vector with missing coordinates. Due to the rigidity of the moving objects, it was noted in [19] that the behavior of observed and missing data is the same and thus they both generate an analogous (frequency) spectral representation. Motivated by this observation, the proposed approach is based on the study of changes in frequency spectra on the initial matrix after missing entries are recovered.

Next we describe an implementation of the max-norm constraint matrix completion procedure, which incorporates the rank estimation approach in [19]. Assume without loss of generality that  $\alpha_0$  is known.

- (1) Given the observed partial matrix  $M_S$ , the initial matrix  $M_{ini}$  is obtained by adding the average of the corresponding column to the missing entries of  $M_S$ . Applying the Fast Fourier Transform (FFT) to the columns of  $M_{ini}$  and taking its modulus, i.e.  $F := |FFT(M_{ini})|$ .
- (2) Set an initial rank  $r = 2$  and an upper bound  $r_{\max}$ . Clearly,  $r_{\max} \leq \min\{d_1, d_2\}$  and it can be computed automatically by adding a criteria for stopping the iteration.

- (3) For the current value of  $r$ , using the computational algorithms given in Section 4 with  $R = \alpha_0 \sqrt{r}$  to solve the max-norm constraint optimization (3.4). The resulting estimated full matrix is denoted by  $\hat{M}_r$ .
- (4) Apply the FFT to  $\hat{M}_r$  as in step 1. Write  $F_r = |FFT(\hat{M}_r)|$  and compute the error  $e(r) = \|F - F_r\|_F$ .
- (5) If  $r < r_{\max}$ , set  $r = r + 1$  and go to step 3.

Finally, let

$$r^* = \arg \min_{r=2, \dots, r_{\max}} e(r)$$

and the corresponding  $\hat{M}_{r^*}$  is the final estimate of  $M_0$ . Clearly, the above procedure can be modified by replacing the rank  $r$  with the max-norm  $R$ . A suitable initial value for the max-norm is  $R = \alpha_0 \sqrt{2}$  and at each iteration, increase  $R = R + \delta$  with a fixed step size  $\delta > 0$ . An upper-bound  $R_{\max}$  could be automatically computed by adding some criteria for stopping the iteration.

## 5 Proofs

We prove the main results, Theorems 3.1 and 3.2, in this section. The proofs of a few key technical lemmas including Lemma 3.1 are also given.

### 5.1 Proof of Theorem 3.1

For simplicity, we write  $\hat{M} = \hat{M}_{\max}$  as long as there is no ambiguity. To begin with, noting that  $\hat{M}$  is optimal and  $M_0$  is feasible for the convex optimization problem (3.4), we thus have the basic inequality

$$\frac{1}{n} \sum_{t=1}^n (Y_{it,jt} - \hat{M}_{it,jt})^2 \leq \frac{1}{n} \sum_{t=1}^n (Y_{it,jt} - (M_0)_{it,jt})^2.$$

This, combined with our model assumption that  $Y_{it,jt} = (M_0)_{it,jt} + \sigma \xi_t$ , yields

$$\frac{1}{n} \sum_{t=1}^n \hat{\Delta}_{it,jt}^2 = \frac{1}{n} \sum_{t=1}^n \{\hat{M}_{it,jt} - (M_0)_{it,jt}\}^2 \leq \frac{2\sigma}{n} \sum_{t=1}^n \xi_t \hat{\Delta}_{it,jt}, \quad (5.1)$$

where  $\hat{\Delta} = \hat{M} - M_0 \in \mathcal{K}(2\alpha, 2R)$  is the error matrix. Then we see that the major challenges in proving Theorem 3.1 consist of two parts, bounding the left-hand side of (5.1) from below in a uniform sense and the right-hand side of (5.1) from above.



**Step 1.** (Upper bound). Recall that  $\{\xi_t\}_{t=1}^n$  is a sequence of  $N(0, 1)$  random variables and  $S = \{(i_1, j_1), \dots, (i_n, j_n)\}$  is drawn i.i.d. according to  $\Pi$  on  $[d_1] \times [d_2]$ , define

$$\hat{\mathcal{R}}_n(\alpha, R) := \sup_{M \in \mathcal{K}(\alpha, R)} \left| \frac{1}{n} \sum_{t=1}^n \xi_t M_{i_t, j_t} \right|. \quad (5.2)$$

Due to Maurey and Pisier [34], we obtain that for any realization of the random training set  $S$  and for any  $\delta > 0$ , with probability at least  $1 - \delta$  over  $\xi = \{\xi_t\}$ ,

$$\begin{aligned} & \sup_{M \in \mathcal{K}(\alpha, R)} \left| \frac{1}{n} \sum_{t=1}^n \xi_t M_{i_t, j_t} \right| \\ & \leq \mathbb{E}_\xi \left[ \sup_{M \in \mathcal{K}(\alpha, R)} \left| \frac{1}{n} \sum_{t=1}^n \xi_t M_{i_t, j_t} \right| \right] + \pi \sqrt{\frac{\log(1/\delta) \max_{M \in \mathcal{K}(\alpha, R)} \sum_{t=1}^n M_{i_t, j_t}^2}{2n^2}} \\ & \leq \mathbb{E}_\xi \left[ \sup_{M \in \mathcal{K}(\alpha, R)} \left| \frac{1}{n} \sum_{t=1}^n \xi_t M_{i_t, j_t} \right| \right] + \pi \alpha \sqrt{\frac{\log(1/\delta)}{2n}}. \end{aligned} \quad (5.3)$$

Thus it remains to estimate the following expectation over the class of matrices  $\mathcal{K}(\alpha, R)$ :

$$\mathcal{R}_n := \mathbb{E}_\xi \left[ \sup_{M \in \mathcal{K}(\alpha, R)} \left| \frac{1}{n} \sum_{t=1}^n \xi_t M_{i_t, j_t} \right| \right].$$

As a direct consequence of (2.4), we have

$$\mathcal{R}_n \leq K_G \cdot R \cdot \mathbb{E}_\xi \left[ \sup_{M \in \mathcal{M}_\pm} \left| \frac{1}{n} \sum_{t=1}^n \xi_t M_{i_t, j_t} \right| \right], \quad (5.4)$$

where  $\mathcal{M}_\pm$  contains rank-one sign matrices with cardinality  $|\mathcal{M}_\pm| = 2^{d-1}$ . For each  $M \in \mathcal{M}_\pm$ ,  $\sum_{t=1}^n \xi_t M_{i_t, j_t}$  is Gaussian with mean zero and variance  $n$ . Then the expectation of this Gaussian maxima can be bounded by

$$\sqrt{2n \log(|\mathcal{M}_\pm|)} \leq \sqrt{2 \log 2} \sqrt{nd}.$$

Since the upper bound is uniform with respect to all realizations of  $S$ , we conclude that with probability at least  $1 - \delta$  over both the random sample  $S$  and the noise  $\xi_t$ ,

$$\hat{\mathcal{R}}_n(\alpha, R) \leq 4 \left( R \sqrt{\frac{d}{n}} + \alpha \sqrt{\frac{\log(1/\delta)}{n}} \right). \quad (5.5)$$

On the other hand, in the case of sub-exponential noise, i.e.  $\{\xi_t\}$  satisfies the assumption (3.8), it follows from (2.4) that

$$\hat{\mathcal{R}}_n(\alpha, R) \leq K_G \cdot R \cdot \sup_{M \in \mathcal{M}_\pm} \left| \frac{1}{n} \sum_{t=1}^n \xi_t M_{i_t, j_t} \right| \quad \text{with} \quad |\mathcal{M}_\pm| = 2^{d-1}.$$

For any realization of the random training set  $S = \{(i_1, j_1), \dots, (i_n, j_n)\}$  and for any  $M \in \mathcal{M}_\pm$  fixed, using Bernstein-type inequality for sub-exponential random variables [45] yields

$$\mathbb{P}\left\{\left|\frac{1}{n}\sum_{t=1}^n \xi_t M_{i_t, j_t}\right| \geq t\right\} \leq 2 \exp\left\{-c \cdot \min\left(\frac{nt^2}{K^2}, \frac{nt}{K}\right)\right\},$$

where  $c > 0$  is an absolute constant. By the union bound, we obtain that for a sample size  $n \geq d$ ,

$$\hat{\mathcal{R}}_n(\alpha, R) \leq CKR\sqrt{\frac{d}{n}} \quad (5.6)$$

holds with probability at least  $1 - e^{-d}$ .

**Step 2.** (Lower bound). For the given sampling distribution  $\Pi$ , define

$$\|M\|_\Pi^2 = \frac{1}{n} \mathbb{E}_{S \sim \Pi}[\|M_S\|_2^2] = \mathbb{E}_{(i_t, j_t) \sim \Pi}[M_{i_t, j_t}^2] = \sum_{k, l} \pi_{kl} M_{kl}^2,$$

where  $M_S = (M_{i_1, j_1}, \dots, M_{i_n, j_n})^T \in \mathbb{R}^n$  for a given training set  $S$ . Then, for any  $\beta \geq 1, \delta > 0$ , consider the following subset

$$\mathcal{C}(\beta, \delta) := \left\{M \in \mathcal{K}(1, \beta) : \|M\|_\Pi^2 \geq \delta\right\}.$$

Here  $\delta$  can be regarded as a tolerance parameter. The goal is to show that there exists some function  $f_\beta$  such that with high probability, the following inequality

$$\frac{1}{n} \|M_S\|_2^2 \geq \frac{1}{2} \|M\|_\Pi^2 - f_\beta(n, d_1, d_2) \quad (5.7)$$

holds for all  $M \in \mathcal{C}(\beta, \delta)$ .

**Proof of (5.7).** Instead, we will prove a stronger result that

$$\left|\frac{1}{n} \|M_S\|_2^2 - \|M\|_\Pi^2\right| \leq \frac{1}{2} \|M\|_\Pi^2 + f_\beta(n, d_1, d_2)$$

for all  $M \in \mathcal{C}(\beta, \delta)$ , with high probability. Following the peeling argument as in [32], for  $\ell = 1, 2, \dots$  and  $\alpha = \frac{3}{2}$ , define a sequence of subsets

$$\mathcal{C}_\ell(\beta, \delta) := \left\{M \in \mathcal{C}(\beta, \delta) : \alpha^{\ell-1} \delta \leq \|M\|_\Pi^2 \leq \alpha^\ell \delta\right\}$$

and for any radius  $D > 0$ , set

$$\mathcal{B}(D) := \left\{M \in \mathcal{C}(\beta, \delta) : \|M\|_\Pi^2 \leq D\right\}. \quad (5.8)$$

Therefore, if there exists some  $M \in \mathcal{C}(\beta, \delta)$  satisfying

$$\left|\frac{1}{n} \|M_S\|_2^2 - \|M\|_\Pi^2\right| > \frac{1}{2} \|M\|_\Pi^2 + f_\beta(n, d_1, d_2),$$

then there corresponds an  $\ell \geq 1$  such that,  $M \in \mathcal{C}_\ell(\beta, \delta) \subset \mathcal{B}(\alpha^\ell \delta)$  and

$$\left| \frac{1}{n} \|M_S\|_2^2 - \|M\|_\Pi^2 \right| > \frac{1}{3} \alpha^\ell \delta + f_\beta(n, d_1, d_2).$$

So the main task is to show that the latter event occurs with high probability. To this end, define the maximum deviation

$$\Delta_D(S) := \sup_{M \in \mathcal{B}(D)} \left| n^{-1} \|M_S\|_2^2 - \|M\|_\Pi^2 \right|. \quad (5.9)$$

The following lemma shows that  $n^{-1} \|M_S\|_2^2$  does not deviate far from its expectation *uniformly* for all  $M \in \mathcal{B}(D)$ .

**Lemma 5.1** (Concentration). *There exists a universal positive constant  $C_1$  such that for any radius  $D > 0$ ,*

$$\mathbb{P} \left\{ \Delta_D(S) > \frac{D}{3} + C_1 \beta \sqrt{\frac{d}{n}} \right\} \leq e^{-nD/10}. \quad (5.10)$$

In view of the above lemma, we can set  $f_\beta(n, d_1, d_2) = C_1 \beta \sqrt{\frac{d}{n}}$  and consider the following sequence of events

$$\mathcal{E}_\ell = \left\{ \Delta_{\alpha^\ell \delta}(S) > \frac{1}{3} \alpha^\ell \delta + f_\beta(n, d_1, d_2) \right\}, \quad \ell = 1, 2, \dots$$

Since  $\mathcal{C}(\beta, \delta) = \cup_{\ell \geq 1} \mathcal{C}_\ell(\beta, \delta)$ , using the union bound we have

$$\begin{aligned} & \mathbb{P} \left\{ \exists M \in \mathcal{C}(\beta, \delta), \text{ s.t. } \left| \frac{1}{n} \|M_S\|_2^2 - \|M\|_\Pi^2 \right| > \frac{1}{2} \|M\|_\Pi^2 + f_\beta(n, d_1, d_2) \right\} \\ & \leq \sum_{\ell \geq 1} \mathbb{P} \left\{ \exists M \in \mathcal{C}_\ell(\beta, \delta), \text{ s.t. } \left| \frac{1}{n} \|M_S\|_2^2 - \|M\|_\Pi^2 \right| > \frac{1}{2} \|M\|_\Pi^2 + f_\beta(n, d_1, d_2) \right\} \\ & \leq \sum_{\ell=1}^{\infty} P(\mathcal{E}_\ell^c) \leq \sum_{\ell=1}^{\infty} \exp(-n\alpha^\ell \delta / 10) \\ & \leq \sum_{\ell=1}^{\infty} \exp\{-\log(\alpha) \ell n \delta / 10\} \leq \frac{\exp(-c_0 n \delta)}{1 - \exp(-c_0 n \delta)} \end{aligned}$$

with  $c_0 = \log(3/2)/10$ , where we used the elementary inequality that

$$\alpha^\ell = \exp\{\ell \log(\alpha)\} \geq \ell \log(\alpha).$$

Consequently, for a sample size  $n \leq d_1 d_2$  satisfying  $\exp(-c_0 n \delta) \leq \frac{1}{2}$ , or equivalently,  $n > \frac{\log 2}{c_0 \delta}$ , we obtain that

$$\frac{1}{n} \|M_S\|_2^2 \geq \frac{1}{2} \|M\|_\Pi^2 - C_1 \beta \sqrt{\frac{d}{n}} \quad \text{for all } M \in \mathcal{C}(\beta, \delta) \quad (5.11)$$

with probability greater than  $1 - 2\exp(-c_0 n \delta)$ .

**Step 3.** Now we combine the results in *Step 1* and *Step 2* to finish the proof. On one hand, it follows from (5.5) that for a sample size  $2 < n \leq d_1 d_2$ ,

$$\frac{1}{n} \sum_{t=1}^n \xi_t \hat{\Delta}(i_t, j_t) \leq \hat{\mathcal{R}}_n(2\alpha, 2R) \leq 8(R + \alpha) \sqrt{\frac{d}{n}}$$

holds with probability at least  $1 - e^{-d}$ . On the other hand, set  $\tilde{\Delta} = \hat{\Delta}/(2\alpha)$ , so that  $\|\tilde{\Delta}\|_\infty \leq 1$  and  $\|\tilde{\Delta}\|_{\max} \leq R/\alpha := \beta$ . Then for any  $t > 0$ , applying (5.11) with  $\delta = \frac{\log(2/t)}{c_0 n}$  yields that for a sample size  $2 < n \leq d_1 d_2$ ,

$$\|\tilde{\Delta}\|_{\Pi}^2 \leq \max \left\{ \delta, \frac{2}{n} \|\tilde{\Delta}_S\|_2^2 + 2\beta C_1 \sqrt{\frac{d}{n}} \right\}$$

with probability at least  $1 - t$ . Above estimates, joint with the basic inequality (5.1) implies the final conclusion (3.5) after a simple rescaling. Similarly, using the upper bound (5.6), instead of (5.5), together with the lower bound (5.11) gives (3.9) in the case of sub-exponential noise. ■

### 5.1.1 Proof of Lemma 5.1

Here we prove the concentration inequality given by Lemma 5.1. The argument is based on techniques of probability in Banach spaces, including symmetrization, contraction inequality and Bousquet's version of Talagrand concentration inequality, and the upper bound (2.6) on the empirical Rademacher complexity of the max-norm ball.

Consider each matrix  $M \in \mathbb{R}^{d_1 \times d_2}$  as a function:  $[d_1] \times [d_2] \rightarrow \mathbb{R}$ , i.e.  $M(k, l) = M_{kl}$ , and rewrite the empirical process of interest as follows:

$$\Delta_D(S) = \sup_{f_M: M \in \mathcal{B}(D)} \left| \frac{1}{n} \sum_{t=1}^n f_M(i_t, j_t) - \mathbb{E}[f_M(i_t, j_t)] \right|, \quad f_M(\cdot) = \{M(\cdot)\}^2.$$

Recall that  $|M_{kl}| \leq \|M\|_\infty \leq 1$  for all pairs  $(k, l)$  and

$$\sup_{M \in \mathcal{B}(D)} \text{Var}[f_M(i_1, j_1)] \leq \sup_{M \in \mathcal{B}(D)} \|M\|_\infty^2 \|M\|_{\Pi}^2 \leq D.$$

We first bound  $\mathbb{E}_{S \sim \Pi}[\Delta_D(S)]$ , then show that  $\Delta_D(S)$  is concentrated around its mean. Now a standard symmetrization argument [25] yields

$$\mathbb{E}_{S \sim \Pi}[\Delta_D(S)] \leq 2\mathbb{E}_{S \sim \Pi} \left\{ \mathbb{E}_\varepsilon \left[ \sup_{M \in \mathcal{B}(D)} \left| \frac{1}{n} \sum_{i=1}^n \varepsilon_i M_{i_t, j_t}^2 \right| \middle| S \right] \right\},$$

where  $\{\varepsilon_i\}_{i=1}^n$  is an i.i.d. Rademacher sequence, independent of  $S$ . Given an index set  $S = \{(i_1, j_1), \dots, (i_n, j_n)\}$ , since  $|M_{i_t, j_t}| \leq 1$ , using Ledoux-Talagrand contraction inequality [25] implies ( $d = d_1 + d_2$ )

$$\begin{aligned} \mathbb{E}_\varepsilon \left[ \sup_{M \in \mathcal{B}(D)} \left| \frac{1}{n} \sum_{i=1}^n \varepsilon_i M_{i_t, j_t}^2 \right| \middle| S \right] &\leq 4\mathbb{E}_\varepsilon \left[ \sup_{M \in \mathcal{B}(D)} \left| \frac{1}{n} \sum_{t=1}^n \varepsilon_t M_{i_t, j_t} \right| \middle| S \right] \\ &\leq 4\mathbb{E}_\varepsilon \left[ \sup_{\|M\|_{\max} \leq \beta} \left| \frac{1}{n} \sum_{t=1}^n \varepsilon_t M_{i_t, j_t} \right| \middle| S \right] \\ &\leq 48\beta \sqrt{\frac{d}{n}}, \end{aligned}$$

where the last step used (2.6). Now that the “worst-case” Rademacher complexity is uniformly bounded, we have

$$\mathbb{E}_{S \sim \Pi}[\Delta_D(S)] \leq 96\beta \sqrt{\frac{d}{n}}. \quad (5.12)$$

Next, using Bousquet’s version of Talagrand concentration inequality for empirical processes indexed by bounded functions implies that for all  $t > 0$ , with probability at least  $1 - e^{-t}$ ,

$$\Delta_D(S) \leq 2 \left[ \mathbb{E}_{S \sim \Pi}[\Delta_D(S)] + \sqrt{\frac{Dt}{n}} + \frac{t}{n} \right].$$

So our conclusion (5.10) follows by taking  $t = \frac{nD}{10}$ . ■

## 5.2 Proof of Theorem 3.2

By construction in Lemma 3.1, set  $\delta = \gamma\alpha\sqrt{d_1 d_2}/2$  and we see that  $\mathcal{M}$  is a  $\delta$ -packing set of  $\mathcal{K}(\alpha, R)$  in the Frobenius norm. Next, a standard argument (e.g. [46, 47]) yields a lower bound on the  $\|\cdot\|_F$ -risk in terms of the error in a multi-way hypothesis testing problem. More concretely,

$$\inf_{\tilde{M}} \max_{M \in \mathcal{K}(\alpha, R)} \mathbb{E} \|\hat{M} - M\|_F^2 \geq \frac{\delta^2}{4} \min_{\tilde{M}} \mathbb{P}(\tilde{M} \neq M^*),$$

where the random variable  $M^* \in \mathbb{R}^{d_1 \times d_2}$  is uniformly distributed over the packing set  $\mathcal{M}$ . Conditional on  $S = \{(i_1, j_1), \dots, (i_n, j_n)\}$ , a variant of Fano’s inequality [11] gives the lower bound

$$\mathbb{P}(\tilde{M} \neq M^* | S) \geq 1 - \frac{\binom{|\mathcal{M}|}{2}^{-1} \sum_{k \neq l} K(M^k \| M^l) + \log 2}{\log |\mathcal{M}|}, \quad (5.13)$$

where  $K(M^k \| M^l)$  is the Kullback-Leibler divergence between distributions  $(Y_S | M^k)$  and  $(Y_S | M^l)$ . For the observation model (3.1) with i.i.d. Gaussian noise, we have

$$K(M^k \| M^l) = \frac{1}{2\sigma^2} \sum_{t=1}^n (M^k - M^l)_{i_t, j_t}^2$$

and

$$\mathbb{E}_S[K(M^k \| M^l)] = \frac{n}{2\sigma^2} \|M^k - M^l\|_\Pi^2, \quad (5.14)$$

where  $\|\cdot\|_\Pi$  denotes the weighted Frobenious norm with respect to  $\Pi$ , that is,

$$\|M\|_\Pi^2 = \sum_{k=1}^{d_1} \sum_{l=1}^{d_2} \pi_{kl} M_{kl}^2, \quad \text{for any } M \in \mathbb{R}^{d_1 \times d_2}.$$

For any two distinct  $M^k, M^l \in \mathcal{M}$ ,  $\|M^k - M^l\|_F^2 \leq 4d_1d_2\gamma^2$ , which together with (5.13), (5.14) and the assumption  $\max_{k,l} \{\pi_{kl}\} \leq \frac{L}{d_1d_2}$  implies

$$\begin{aligned} \mathbb{P}(\tilde{M} \neq M^*) &\geq 1 - \frac{\binom{|\mathcal{M}|}{2}^{-1} \sum_{k \neq l} \mathbb{E}_S[K(M^k \| M^l)] + \log 2}{\log |\mathcal{M}|} \\ &\geq 1 - \frac{\frac{32L\gamma^4\alpha^2n}{\sigma^2} + 12\gamma^2}{r(d_1 \vee d_2)} \geq 1 - \frac{32L\gamma^4\alpha^2n}{\sigma^2 r(d_1 \vee d_2)} - \frac{12}{r(d_1 \vee d_2)} \geq \frac{1}{2}, \end{aligned}$$

provided that  $r(d_1 \vee d_2) \geq 48$  and  $\gamma^4 \leq \frac{\sigma^2}{128L\alpha^2} \frac{r(d_1 \vee d_2)}{n}$ . If  $\frac{\sigma^2}{128L\alpha^2} \frac{r(d_1 \vee d_2)}{n} > 1$ , then we choose  $\gamma = 1$  so that

$$\inf_{\hat{M}} \max_{M \in \mathcal{K}(\alpha, r)} \frac{1}{d_1d_2} \mathbb{E} \|\hat{M} - M\|_F^2 \geq \frac{\alpha^2}{16}.$$

Otherwise, as long as the parameters  $(n, d_1, d_2, \alpha, R)$  satisfies (3.11), setting

$$\gamma^2 = \frac{\sigma}{8\sqrt{2}\alpha} \sqrt{\frac{r(d_1 \vee d_2)}{nL}}$$

yields

$$\inf_{\hat{M}} \max_{M \in \mathbb{B}_{\max}(R)} \frac{1}{d_1d_2} \mathbb{E} \|\hat{M} - M\|_F^2 \geq \frac{\sigma\alpha}{128\sqrt{2}} \sqrt{\frac{r(d_1 \vee d_2)}{nL}} \geq \frac{\sigma R}{256} \sqrt{\frac{d}{nL}},$$

as desired.  $\blacksquare$

### 5.3 Proof of Lemma 3.1

We proceed via the probabilistic method. Assume without loss of generality that  $d_2 \geq d_1$ . Let  $N = \exp(\frac{rd_2}{16\gamma^2})$ ,  $B = \frac{r}{\gamma^2}$ , and for each  $i = 1, \dots, N$ , we draw a random matrix  $M^i \in \mathbb{R}^{d_1 \times d_2}$  as follows: the matrix  $M^i$  consists of i.i.d. blocks of dimensions  $B \times d_2$ , stacked from top to bottom, with the entries of the first block being i.i.d. symmetric random variables taking values  $\pm\alpha\gamma$ , such that

$$M_{kl}^i := M_{k'l}^i, \quad k' = k(\bmod B) + 1.$$

Next we show that above random procedure succeeds in generating a set having all desired properties, with non-zero probability. For  $1 \leq i \leq N$ , it is easy to see that

$$\|M^i\|_\infty = \alpha\gamma \leq \alpha, \quad \frac{1}{d_1 d_2} \|M^i\|_F^2 = \alpha^2 \gamma^2$$

and since  $\text{rank}(M^i) \leq B$ ,

$$\|M^i\|_{\max} \leq \sqrt{B} \|M^i\|_\infty = \sqrt{\frac{r}{\gamma^2}} \alpha \gamma = \alpha \sqrt{r} = R.$$

Thus  $M^i \in \mathcal{K}(\alpha, R)$ , and it remains to show that the set  $\{M^i\}_{i=1}^N$  satisfies property (ii). In fact, for any pair  $1 \leq i \neq j \leq N$ ,

$$\|M^i - M^j\|_F^2 = \sum_{k,l} (M_{kl}^i - M_{kl}^j)^2 \geq \left\lfloor \frac{d_1}{B} \right\rfloor \sum_{k=1}^B \sum_{j=1}^{d_2} (M_{kl}^i - M_{kl}^j)^2 = 4\alpha^2 \gamma^2 \left\lfloor \frac{d_1}{B} \right\rfloor \sum_{k=1}^B \sum_{j=1}^{d_2} \delta_{kl},$$

where  $\delta_{kl}$  are independent 0/1 Bernoulli random variables with mean 1/2. Using the Hoeffding's inequality gives

$$\mathbb{P} \left\{ \sum_{k=1}^B \sum_{j=1}^{d_2} \delta_{kl} \geq \frac{d_2 B}{4} \right\} \leq \exp(-d_2 B/8).$$

Since there are less than  $N^2/2$  such index pairs in total, above inequality, together with a union bound implies that with probability at least  $1 - \frac{N^2}{2} \exp(-d_2 B/8) \geq 1/2$ ,

$$\|M^i - M^j\|_F^2 > \alpha^2 \gamma^2 \left\lfloor \frac{d_1}{B} \right\rfloor d_2 B \geq \frac{\alpha^2 \gamma^2 d_1 d_2}{2}, \quad \text{for all } i \neq j.$$

This completes the proof of Lemma 3.1.  $\blacksquare$

## References

- [1] ARGYRIOU, A., EVGENIOU, T. and Pontil, M. (2008). Convex multi-task feature learning. *Mach. Learn.* **73** 243-272.
- [2] BARTLETT, P. and MENDELSON, S. (2002). Rademacher and Gaussian complexities: Risk bounds and structural results. *J. Mach. Learn. Res.* **3** 463-482.
- [3] BISWAS, P., Lian, T.-C., WANG, T.-C. and Ye, Y. (2006). Semidefinite programming based algorithms for sensor network localization. *ACM Trans. Sen. Netw.* **2** 188-220.
- [4] BURER, S. and CHOI, C. (2006). Computational enhancements in low-rank semidefinite programming. *Optimization Methods and Software* **21** 493-512.

- [5] CHISTOV, A.L. and GRIGORIEV, D.Y. (1984). Complexity of quantifier elimination in the theory of algebraically closed fields. In *Proceedings of the 11th Symposium on Mathematical Foundations of Computer Science*, volume 176 of *Lecture Notes in Computer Science*, pages 17-31. Springer Verlag.
- [6] CANDÈS, E. and PLAN, Y. (2010). Matrix completion with noise. *Proc. IEEE* **98** 925-936.
- [7] CANDÈS, E. and PLAN, Y. (2011). Tight oracle bounds for low-rank matrix recovery from a minimal number of random measurements. *IEEE Trans. Inform. Theory* **57** 2342-2359.
- [8] CANDÈS, E. and RECHT, B. (2009). Exact matrix completion via convex optimization. *Found. Comput. Math.* **9** 717-772.
- [9] CANDÈS, E. and TAO, T. (2010). The power of convex relaxations: Near-optimal matrix completion. *IEEE Trans. Inform. Theory* **56** 2053-2080.
- [10] CHEN, P. and SUTER, D. (2004). Recovering the missing components in a large noisy low-rank matrix: application to SFM. *IEEE Trans. Pattern Anal. Mach. Intell.* **26** 1051-1063.
- [11] COVER, T.M. and THOMAS, J.A. (1991). *Elements of Information Theory*. John Wiley and Sons, New York.
- [12] DAVENPORT, M.A., PLAN, Y., VAN DEN BERG, E. and WOOTTERS, M. (2012). 1-bit matrix completion. *arXiv:1209.3672*.
- [13] FOYGEL, R., SALAKHUTDINOV, R., SHAMIR, R. and SREBRO, N. (2011). Learning with the weighted trace-norm under arbitrary sampling distributions. *Advances in Neural Information Processing Systems (NIPS)*, **24**.
- [14] FOYGEL, R. and SREBRO, N. (2011). Concentration-based guarantees for low-rank matrix reconstruction. *24th Annual Conference on Learning Theory (COLT)*.
- [15] GOLDBERG, D., NICHOLS, D., OKI, B.M. and TERRY, D. (1992). Using collaborative filtering to weave an information tapestry. *Comm. ACM* **35** 61-70.
- [16] GREEN, P. and WIND, Y. (1973). *Multivariate decisions in marketing: A measurement approach*. Dryden, Hinsdale, IL.
- [17] GROSS, D. (2011). Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. Inform. Theory* **57** 1548-1566.



- [18] JAMESON, G.J.O. (1987). *Summing and Nuclear Norms in Banach Space Theory*. Number 8 in London Mathematical Society Student Texts. Cambridge University Press, Cambridge, UK.
- [19] JULIÀ, C., SAPPÀ, A.D., LUMBRERAS, F., SERRAT, J. and LÓPEZ, A. (2011). Rank estimation in missing data matrix problems. *J. Math. imaging Vis.* **39** 140-160.
- [20] KESHAVAN, R., MONTANARI, A. and OH, S. (2010). Matrix completion from noisy entries. *J. Mach. Learn. Res.* **11** 2057-2078.
- [21] KLOPP, O. (2012). Noisy low-rank matrix completion with general sampling distribution. *arXiv:1203.0108*.
- [22] KLOPP, O. (2011). Rank penalized estimators for high-dimensional matrices. *Electron. J. Stat.* **5** 1161-1183.
- [23] KOLTCHINSKII, V. (2011). Von Neumann entropy penalization and low-rank matrix estimation. *Ann. Statist.* **39** 2936-2973.
- [24] KOLTCHINSKII, V., LOUNICI, K. and TSYBAKOV, A.B. (2011). Nuclear norm penalization and optimal rates for noisy low rank matrix completion. *Ann. Statist.* **39** 2302-2329.
- [25] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces: Isoperimetry and Processes*. Springer-Verlag, New York, NY.
- [26] LEE, J., RECHT, B., SALAKHUTDINOV, R., Srebro, N. and TROPP, J. (2010). Practical large-scale optimization for max-norm regularization. *Advances in Neural Information Processing Systems*, **23**.
- [27] LIN, Z., GANESH, A., WRIGHT, J., WU, L., Chen, M. and MA, Y. (2009). Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix. In *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*.
- [28] LINIAL, N., MENDELSON, S., SCHECHTMAN, G. and SHRAIBMAN, A. (2004). Complexity measures of sign measures. *Combinatorica* **27** 439-463.
- [29] LIU, Z. and VANDENBERGHE, L. (2009). Interior-point method for nuclear norm approximation with application to system identification. *SIAM J. Matrix Analysis and Applications*, **31** 1235-1256.
- [30] MAZUMBER, R., HASTIE, T. and TIBSHIRANI, R. (2010). Spectral regularization algorithms for learning large incomplete matrices. *J. Mach. Learn. Res.* **11** 2287-2322.

- [31] NEGAHBAN, S. and WAINWRIGHT, M.J. (2011). Estimation of (near) low-rank matrices with noise and high-dimensional scaling. *Ann. Statist.* **39** 1069-1097.
- [32] NEGAHBAN, S. and WAINWRIGHT, M.J. (2012). Restricted strong convexity and weighted matrix completion: Optimal bounds with noise. *J. Mach. Learn. Res.* **13** 1665-1697.
- [33] NESTEROV, Y. (2007). Gradient methods for minimizing composite objective function. *Technical Report - CORE - Universite Catholique de Louvain*.
- [34] PISIER, G. (1989). *The volume of convex bodies and Banach space geometry*. Cambridge University Press, Cambridge.
- [35] RECHT, B. (2011). A simpler approach to matrix completion. *J. Mach. Learn. Res.* **12** 3413-3430.
- [36] ROHDE, A. and TSYBAKOV, A.B. (2011). Estimation of high-dimensional low-rank matrices. *Ann. Statist.* **39** 887-930.
- [37] SALAKHUTDINOV, R. and SREBRO, N. (2010). Collaborative filtering in a non-uniform world: Learning with the weighted trace norm. *Advances in Neural Information Processing Systems (NIPS)*, **23**.
- [38] SINGER, A. (2008). A remark on global positioning from local distances. *Proc. Natl. Acad. Sci.* **105** 9507-9511.
- [39] SINGER, A. and CUCURINGU, M. (2010). Uniqueness of low-rank matrix completion by rigidity theory. *SIAM J. Matrix Analysis and Applications*, **31** 1621-1641.
- [40] SREBRO, N., RENNIE, J. and JAAKKOLA, T. (2004). Maximum-margin matrix factorization. In *Advances in Neural Information Processing Systems* **17** (L. Saul, Y. Weiss and L. Bottou, eds.) 1329-1336. MIT Press, Cambridge, MA.
- [41] SREBRO, N. and SHRAIBMAN, A. (2005). Rank, trace-norm and max-norm. In *Learning Theory, Proceedings of COLT-2005. Lecture Notes in Comput. Sci.* **3559** 545-560. Springer, Berlin.
- [42] SREBRO, N., SRIDHARAN, K. and TEWARI, A. (2010). Optimistic Rates for Learning with a Smooth Loss. *arXiv:1009.3896v2*.
- [43] TOMASI, C. and KANADE, T. (2012). Shape and motion from image streams under orthography: a factorization method. *Int. J. Comput. Vis.* **9** 137-154.

- [44] TROPP, J.A. (2012). User-friendly tail bounds for sums of random matrices. *Found. Comput. Math.* **12** 389-434.
- [45] VERSHYNIN, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing: Theory and Applications*, Y. Eldar and G. Kutyniok, Eds. Cambridge University Press.
- [46] YANG, Y. and BARRON, A. (1999). Information-theoretic determination of minimax rates of convergence. *Ann. Statist.* **27** 1564-1599.
- [47] YU, B. (1997). Assouad, Fano, and Le Cam. *Festschrift for Lucien Le Cam*. D. Pollard, E. Torgersen, and G. Yang (eds), 423-435, Springer-Verlag.